

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 30, 2002, 17:23:45 ; Search time 24.61 Seconds  
(without alignments)  
787.299 Million cell updates/sec

Title: US-09-730-617-4  
Perfect score: 568  
Sequence: 1 MFGSLHFLALAGVPLSM.....LSRAPQIQYRLVQIOLQK 112

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTRMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp Vertebrate:\*  
14: sp Unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriophage:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	535	94.2	121	4	096H5
2	483	85.0	154	4	096A06
3	399	70.2	121	11	09CR53
4	102	18.0	157	13	091829
5	92.5	16.3	148	4	09BSV7
6	85	15.0	119	13	090253
7	76.5	13.5	119	13	090254
8	75	13.2	771	4	096C12
9	74.5	13.1	243	4	09EV74
10	74	13.0	297	16	09A589
11	74	13.0	577	16	09PG72
12	73.5	12.9	412	10	09AK19
13	73.5	12.9	1902	4	014122
14	71.5	12.6	1339	4	075276
15	71.5	12.6	3638	4	015142
16	71.5	12.6	4292	4	015141

17	71.5	12.6	4302	4	015140	015140 homo sapien
18	70.5	12.4	505	16	096CG8	096CG8 rhizobium 1
19	70	12.3	218	2	03571	03571 rhodobacter
20	70	12.3	342	13	09Y150	09Y150 gallus gall
21	69.5	12.2	732	4	09H7P8	09H7P8 homo sapien
22	69.5	12.2	755	5	021119	021119 caenorhabdi
23	69.5	12.2	814	5	0965Y0	0965Y0 caenorhabdi
24	69.5	12.2	1166	4	09ULK3	09ULK3 homo sapien
25	69.5	12.2	1504	10	09LX06	09LX06 arabidopsis
26	69	12.1	394	16	09AAG9	09AAG9 caulobacter
27	69	12.1	474	13	073906	073906 gallus gall
28	69	12.1	756	16	053874	053874 mycobacteri
29	69	12.1	961	2	09AK84	09AK84 streptomyce
30	68.5	12.1	403	10	09M1B4	09M1B4 arabidopsis
31	68.5	12.1	726	6	09XAP0	09XAP0 streptomyce
32	68.5	12.1	745	2	029517	029517 oryctolagus
33	68.5	12.1	756	11	09Z1B4	09Z1B4 mus musculu
34	68.5	12.1	1400	3	09HFK8	09HFK8 neurospora
35	67.5	11.9	292	11	09DAF2	09DAF2 mus musculu
36	67.5	11.9	435	10	0942W8	0942W8 oryza sativ
37	67.5	11.9	558	10	09SWV4	09SWV4 lycopersico
38	67	11.8	186	16	0914T8	0914T8 pseudomonas
39	67	11.8	441	16	055718	055718 synechocyst
40	67	11.8	497	16	09RZV7	09RZV7 delnoccocus
41	67	11.8	698	16	09BR81	09BR81 mycoplasma
42	67	11.8	15281	3	009164	009164 tolypocladl
43	66.5	11.7	323	4	09H936	09H936 homo sapien
44	66.5	11.7	347	16	09KR79	09KR79 delnoccocus
45	66.5	11.7	371	16	09B9W5	09B9W5 rhizobium 1

## ALIGNMENTS

RESULT 1  
ID 096H5 PRELIMINARY; PRT; 121 AA.  
AC 096H5;  
DT 01-DEC-2001 (TRMBLrel. 19, Created)  
DT 01-DEC-2001 (TRMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TRMBLrel. 19, Last annotation update)  
DE NEUROMEDIN B.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN, AND GLOBLASTOMA WITH EGFR AMPLIFICATION;  
RA Strausberg R;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC008603; AAI08603.1;  
SQ SEQUENCE 121 AA; 13255 MW; 17FA0615486F03FB CRC64;

Query Match 94.2% Score 535; DB 4; Length 121;  
Best Local Similarity 95.5% Pred. No. 2.2e-53;  
Matches 107; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
OY 1 MFGSLHFLALAGVPLSM...PEPRSRASKIRVHSRCKLMAIGHFMGKKSLSPSSPSPL 60  
DB 10 MFGSLHFLALAGVPLSM...PEPRSRASKIRVHSRCKLMAIGHFMGKKSLSPSSPSPL 69  
OY 61 GTAAHTSLRDORLSDHDLGITLLKALGVLSRAPQIQYRLVQIOLQK 112  
DB 70 GTAAHTSLRDORLSDHDLGITLLKALGVLSRAPQIQYRLVQIOLQK 121  
RESULT 2  
ID 096A06 PRELIMINARY; PRT; 154 AA.  
AC 096A06;  
DT 01-DEC-2001 (TRMBLrel. 19, Created)

VS  
12/6/99

DT 01-DEC-2001 (Tremblrel. 19, last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE SIMILAR TO NEUROMEDIN B.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=OVARY, AND ADENOCARCINOMA;  
 RA Strausberg R.;  
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC007431; AA07431.1;  
 DR EMBL: BC007407; AA07407.1;  
 SO SEQUENCE 154 AA; 16313 MW; 3FECAS687F548044 CRC64;

Query Match 85.0%; Score 483; DB 4; Length 154;  
 Best Local Similarity 95.0%; Pred. No. 2.5e-47;  
 Matches 96; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 MEGSLHFLALAGVPLSMDEPPRSRASKIRVHSGKLMAGHFMGKKSLEPSSPPL 60  
 DB 10 MEGSLLEALLAGVAPLSPMDPEPRSRASKIRVHSGKLMATGHPMGKKSLEPSSPPL 69  
 OY 61 GTAPHTSLRDRLQSLSHDLGILLKALGVSLSPAPQIQ 101  
 DB 70 GTAPHTSLRDRLQSLSHDLGILLKALGVSLSPAPQIQ 110

RESULT 3  
 O9CR53 PRELIMINARY; PRT; 121 AA.  
 AC O9CR53;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, last sequence update)  
 DE 3110023K12RIK. PROTEIN.  
 GN 3110023K12RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCBL\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD, AND EMBRYO;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Hono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Flischiemann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Queckenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guenichon S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,  
 RA Yashuhara-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK014077; BAB29144.1;  
 DR EMBL: AK011929; BAB27922.1;  
 DR MGD: MGI:1915289; 3110023K12RIK.  
 OR InterPro: IPR000874; Bombsin.  
 DR Pfam: PF02044; Bombsin; 1.  
 DR PROSITE: PS00257; BOMBESIN; 1.

SO SEQUENCE 121 AA; 13571 MW; 1EE94C963D309D31 CRC64;

Query Match 70.2%; Score 399; DB 11; Length 121;  
 Best Local Similarity 75.0%; Pred. No. 7.2e-38;  
 Matches 81; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

OY 5 LHLFALLAGVPLSMDEPPRSRASKIRVHSGKLMAGHFMGKKSLEPSSPPLGTAP 64  
 DB 14 LHLFALLAGVAPLSPMDPEPRSRASKIRVHSGKLMATGHPMGKKSLEPSSPPLVGTAP 73  
 OY 65 HTSLRDRLQSLSHDLGILLKALGVSLSPAPQIQIRLLVQLOK 112  
 DB 74 PHTPRDRLQSLSHDLGILLKALGVSLSPAPQIQIRLLVQLOK 121

RESULT 4  
 O91829 PRELIMINARY; PRT; 157 AA.  
 AC O91829;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DE 01-DEC-2001 (Tremblrel. 19, last annotation update)  
 DE GASTRIN-RELEASING PEPTIDE PRECURSOR.  
 GN GRP.  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Cyprinidae; Carassius.  
 NC NCBL\_TaxID=7957;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20336817; PubMed=10876046;  
 RA Volkoff H., Peyon P., Lin X., Peter R.E.;  
 RT "Molecular cloning and expression of cDNA encoding a brain  
 bombesin/gastrin-releasing peptide-like peptide in goldfish.";  
 RL Peptides 21:639-648(2000).  
 DR EMBL: AF111028; AAF82387.1;  
 DR InterPro: IPR000874; Bombsin.  
 DR Pfam: PF02044; Bombsin; 1.  
 DR PROSITE: PS00257; BOMBESIN; 1.  
 KW SIGNAL.  
 FT SIGNAL 1 27 POTENTIAL.  
 FT CHAIN 28 157 GASTRIN-RELEASING PEPTIDE.  
 SO SEQUENCE 157 AA; 18201 MW; DD927EE19FD43117 CRC64;

Query Match 18.0%; Score 102; DB 13; Length 157;  
 Best Local Similarity 31.9%; Pred. No. 0.00077;  
 Matches 29; Conservative 13; Mismatches 37; Indels 12; Gaps 3;

OY 33 RVHSGKLMAGHFMGKKSLEPSSPPLGTAPHTSLRDRLQSLSHDLGILLKALGVSLSPAPQIQIRLLVQLOK 112  
 DB 35 KVPYRGNHAGVHMGKSTDEQVAPREPED-GDTSMTTRDQDLKLYNKHLLPLLR 93  
 OY 82 ILLKALGVSLSPAPQIQIRLLVQLOK 112  
 DB 94 VLMGRMAPESMLGTEEDTHTKLQHLIEE 124

RESULT 5  
 O9BSY7 PRELIMINARY; PRT; 148 AA.  
 AC O9BSY7;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, last sequence update)  
 DE 01-OCT-2001 (Tremblrel. 18, last annotation update)  
 DE GASTRIN-RELEASING PEPTIDE.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC TISSUE-LUNG CARCINOMA;  
 RA Strausberg R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC004488; AAH04488.1; -  
 DR InterPro: IPR000874; Bombesin.  
 DR Pfam: PF02044; Bombesin; 1.  
 DR PROSITE: PS00257; BOMBESIN; 1.  
 SO SEQUENCE 148 AA; 16213 MW; 6A9ED9E2748255BD CRC64;

Query Match 16.3%; Score 92.5; DB 4; Length 148;  
 Best Local Similarity 33.6%; Pred. No. 0.0088;  
 Matches 39; Conservative 13; Mismatches 27; Indels 37; Gaps 7;

QY 1 MGSLLHFALLAGVPLSMDLPEPRSRASKI-----RVHSRCKLMAIGHMGKKS 51  
 DB 1 MGRRELPLVLLAL-VLCIA-----PRGRAVPLPAGGGTGLTKMYPGCHMAVAGHMGKKS 54  
 QY 52 L-EPSSPPLGTAPHTSLRDRL-----QLSHDLGILLIK-----KALG 90  
 DB 55 TGESSSVSENG-----SLKQQLREYTRMEEARNLGLIAKENRNHPQPRALG 105

RESULT 6  
 ID 090253 PRELIMINARY; PRT; 119 AA.

AC 090253;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE PHE-13 BOMBESIN PREPROHORMONE.  
 OS Bombina orientalis (Oriental fire-bellied toad).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Archaeobatrachia; Bombinatoridae; Bombina.  
 OX NCBI\_TaxID=8346;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96205965; PubMed=8631814;  
 RA Nagalla S.R., Barry B.J., Fallick A.M., Gibson B.W., Taylor J.E.,  
 RA Dong J.Z., Spindel E.R.;  
 RT "There are three distinct forms of bombesin. Identification of  
 RT [leu13]bombesin, [phe13]bombesin, and [ser3,Arg10,Phe13]bombesin in  
 RT the frog Bombina orientalis."  
 RL J. Biol. Chem. 271:7731-7737(1996).  
 DR EMBL: U049450; AAC59784.1; -  
 DR InterPro: IPR000874; Bombesin.  
 DR Pfam: PF02044; Bombesin; 1.  
 DR PROSITE: PS00257; BOMBESIN; 1.  
 FT CHAIN 45 59 PHE-13 BOMBESIN.  
 SO SEQUENCE 119 AA; 13888 MW; 3C9BD367441E99A CRC64;

Query Match 15.0%; Score 85; DB 13; Length 119;  
 Best Local Similarity 27.0%; Pred. No. 0.049;  
 Matches 31; Conservative 19; Mismatches 49; Indels 16; Gaps 5;

QY 3 GSLLHFALLAGVPLS--MDLPEPRSRASKIRVHSR-GKIMAGHMGKKSLEPSSPSP 59  
 DB 13 GFLFH--LLIFSFALSSCHEFVEDPNNQGLSLQOGLGOMAWGHMGKKSLODTDFEE 70  
 QY 60 LGTAPHTSLRDRLQLSHDLGILLIKALGVSLSRAPQTO--YRRLVQIIOIK 112  
 DB 71 MESFAKRNVENMR-----AALLQDQNRABESERELRNVLAVRNILQYLLK 116

RESULT 7  
 ID 090254 PRELIMINARY; PRT; 119 AA.  
 AC 090254;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

DE SAP BOMBESIN PREPROHORMONE.  
 OS Bombina orientalis (Oriental fire-bellied toad).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Archaeobatrachia; Bombinatoridae; Bombina.  
 OX NCBI\_TaxID=8346;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96205965; PubMed=8631814;  
 RA Nagalla S.R., Barry B.J., Fallick A.M., Gibson B.W., Taylor J.E.,  
 RA Dong J.Z., Spindel E.R.;  
 RT "There are three distinct forms of bombesin. Identification of  
 RT [leu13]bombesin, [phe13]bombesin, and [ser3,Arg10,Phe13]bombesin in  
 RT the frog Bombina orientalis."  
 RL J. Biol. Chem. 271:7731-7737(1996).  
 DR EMBL: U049451; AAC59785.1; -  
 DR InterPro: IPR000874; Bombesin.  
 DR Pfam: PF02044; Bombesin; 1.  
 DR PROSITE: PS00257; BOMBESIN; UNKNOWN\_1.  
 FT CHAIN 45 59 SAP BOMBESIN.  
 SO SEQUENCE 119 AA; 13745 MW; 7F95C34EDEF8702 CRC64;

Query Match 13.5%; Score 76.5; DB 13; Length 119;  
 Best Local Similarity 30.3%; Pred. No. 0.46;  
 Matches 23; Conservative 13; Mismatches 35; Indels 5; Gaps 2;

QY 3 GSLLHFALLAGVPLSMDL---PEPRSRASKIRVHSRCKLMAIGHMGKKSLEPSSPSP 59  
 DB 13 GFLFH--LLIFSFALSSCHEFVEDPNNQGLSLQOGLGOMAWGHMGKKSLODTNFOE 70  
 QY 60 LGTAPHTSLRDRLQL 75  
 DB 71 MESFAKRNVENMRAL 86

RESULT 8  
 ID 096C12 PRELIMINARY; PRT; 771 AA.

AC 096C12;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HYPONETICAL 81.4 KDA PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SKIN, MELANOMA, AND HIGH MDR.  
 RA Strausberg R.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC014945; AAH14945.1; -  
 DR Hypothetical protein.  
 SO SEQUENCE 771 AA; 81368 MW; 98E3B052CD34191B CRC64;

Query Match 13.2%; Score 75; DB 4; Length 771;  
 Best Local Similarity 26.7%; Pred. No. 5.9;  
 Matches 32; Conservative 16; Mismatches 36; Indels 36; Gaps 6;

QY 2 FGSLLHFALLAGVPLSMDLPEPRSRASKIRVSRGKLMAG-HFMGKKSLEPSSPSP 60  
 DB 423 YGALLLAWLVLYGAPDDW--PAPRAPF--LISKR--ELGSRILQNLTVQAESFPGV 475  
 QY 61 GTAPHTSLRD-----QRLQSHDLGILLIKALGVSLSRAP 98  
 DB 476 GALTHTLLSGSPEDRVACALTPFTCRKPSLMRLLLEGGRLULL-----AALTTPAP 529

RESULT 9  
 ID 09EV74 PRELIMINARY; PRT; 243 AA.



DR SMART; SM00400; znf\_CHCC; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 577 AA; 65661 MW; 805C5ABAA08D9015 CRC64;

Query Match 13.0%; Score 74; DB 16; Length 577;  
 Best Local Similarity 27.8%; Pred. No. 5.5;  
 Matches 27; Conservative 12; Mismatches 36; Indels 22; Gaps 4;

QY 9 ALLAGVPTSLDPE-----PRSRASKIRVHSRGLMAGHFGKKSLEPSSPLGT 62  
 DB 385 ARLEARARPLIMIPGAFNDLMRQ-----LAHLTYVIAHP-GPAPAT 429

QY 63 A-PHTSLRQRLSHDLGILKLAGVLSRPAP 98  
 DB 430 SRPATRLPAOKRNLIHTTIIILQPSLMTLNTPTP 466

RESULT 12  
 ID 094K19 PRELIMINARY; PRT; 412 AA.

DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DE MALE FERTILITY PROTEIN.

OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoideae; Andropogoneae; Zea.

NCBI\_TaxID=4577;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. B73;  
 RA Fox T.W., Trimmel M.R., Albertsen M.C.;  
 RT "Cloning of Ms45, a gene required for male fertility from Zea mays."  
 RL EMBL: AF260356; AAK52489.1; -  
 DR EMBL: AF260356; AAK52489.1; -  
 SQ SEQUENCE 412 AA; 46537 MW; C4781406949E2C68 CRC64;

Query Match 12.9%; Score 73.5; DB 10; Length 412;  
 Best Local Similarity 29.2%; Pred. No. 4.2;  
 Matches 28; Conservative 11; Mismatches 24; Indels 33; Gaps 5;

QY 1 MFGSLHFAALLA---GVPL-----SWDLPEPRSRASKIRVHSR 37  
 DB 21 LFPAALALALVADPFGSLPLAEVDYRPHKELAPYGEVMSG---PRDNASRLR--R 73

QY 38 GKLMAGHFGKKSLEPSSPLGTAPHTSLRDQRL 73  
 DB 74 GRLEFVGVEFGPESTIEFDLQ---GRGYAGLADGRV 106

RESULT 13  
 ID 014122 PRELIMINARY; PRT; 1902 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE DNA-BINDING PROTEIN (MBP-1) (FRAGMENT).

OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.  
 NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=90205817; PubMed=2108316;  
 RA Baldwin A.S., Leclair K.P., Singh H., Sharp P.A.;  
 RT "A large protein containing zinc finger domains binds to related  
 sequence elements in the enhancers of the class I major

RT histocompatibility complex and kappa immunoglobulin genes."  
 RL Mol. Cell. Biol. 10:1406-1414 (1990).

DR EMBL: M32019; AA17534.1; -  
 DR HSSP: P15822; 1BBO.  
 DR InterPro: IPR000822; znf-C2H2.  
 DR Pfam: PF00096; znf-C2H2; 3.  
 DR SMART: SM00355; znf-C2H2; 2.  
 DR PROSITE: PS00028; ZINC\_FINGER\_C2H2\_1; 2.  
 DR PROSITE: PS00157; ZINC\_FINGER\_C2H2\_2; 2.  
 KW DNA-binding; Metal-binding; Zinc-finger.  
 FT NON\_TER 1  
 SQ SEQUENCE 1902 AA; 207457 MW; 35BBF0D961E6ED20 CRC64;

Query Match 12.9%; Score 73.5; DB 4; Length 1902;  
 Best Local Similarity 38.4%; Pred. No. 25;  
 Matches 28; Conservative 4; Mismatches 24; Indels 17; Gaps 3;

QY 17 PLSWDLPEPRSRASKI-----RVHSRGLMAGHFGKKSLEPSSPLGTAPHTSLRD 70  
 DB 329 PMSDFKPEPRSRASPVSGFELNRTGKSGSLKVT-----ISQESHPSRDSHPH----- 378

QY 71 QRLQSHDLGIL 83  
 DB 379 -QLALSDALNGEL 390

RESULT 14  
 ID 075276 PRELIMINARY; PRT; 1339 AA.

DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE PKD1 (FRAGMENT).

GN PKD1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.  
 NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.  
 RA Rieke D.O., Bruce D., Mundt M., Doggett N., Munk C., Saunders E.,  
 RA Robinson D., Jones M., Buckingham J., Chasteen L., Thompson S.,  
 RA Goodwin L., Bryant J., Tesmer J., Melnick L., Longmire J., White S.,  
 RA Ueng S., Tatum O., Campbell C., Fawcett J., Maltbie M., Mistra M.,  
 RA Deaven L.;  
 RT "Sequencing of Human Chromosome 16p13.3";  
 RL submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

RN 12  
 RP SEQUENCE FROM N.A.  
 RA Rieke D.O.;  
 RT "Large Scale Sequence Analysis and Annotation with the Sequence  
 Comparison Analysis (SCAN) System";  
 RL submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AC005600; AAC34211.1; -  
 DR InterPro: IPR001024; LH2.  
 DR InterPro: IPR000434; PKD\_1.  
 DR InterPro: IPR000203; PKD\_Cys-rich.  
 DR Pfam: PF01825; GPS; 1.  
 DR Pfam: PF01477; PLAT; 1.  
 DR PRINTS: PR00500; POLYCYSTIN1.  
 DR SMART: SM00303; GPS; 1.  
 DR SMART: SM00308; LH2; 1.  
 FT NON\_TER 1

SO SEQUENCE 1339 AA; 147095 MW; 8C30F58A3ED2B891 CRC64;

Query Match 12.6%; Score 71.5; DB 4; Length 1339;  
 Best Local Similarity 26.0%; Pred. No. 28;  
 Matches 34; Conservative 17; Mismatches 51; Indels 29; Gaps 5;

QY 3 GSLHFAALLAGVP-----LSWDLPEPRSRASKIRV--HSRGLMAGHFGKKSLE 53

```

Db 169 GTTAVHGIMLYGVDSRSGRHLDGDRAFHRNSLDIFRIATPHSLGSKVRWHDNKGCL- 227
OY 54 PSSPPLCTAPHTSLRDOR-----LQSHDLGLGILLKKAIGVSLSRPAPQIO 101
Db 228 ----SPANFLOHVIVRDQGTARSAFPLYNDWLSVETENGGLVEKVLAA--DALLR 280
OY 102 YRRLVQIOLK 112
Db 281 FRRLVAVELOR 291

RESULT 15
O15142 PRELIMINARY: PRT: 3638 AA.
AC O15142:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE POLYCYSTIC KIDNEY DISEASE-ASSOCIATED PROTEIN.
GN PKD1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94273192; PubMed=8004675;
RA Ward C.J., Peral B., Hughes J., Thomas S., Gamble V., McCarthy A.B.,
RA Sloane-Stanley J., Buckle V.J., Kearney L., Higgs D.R.,
RA Ratcliffe P.D., Harris P.C., Koelisema J.H., Spruit L.L., Saris J.J.,
RA Dawerise H.G., Peters D.J.M., Breuning M.H., Nellist M.,
RA Brook-Carter P.T., Maheshwar M.M., Cordeiro I., Santos H., Cabral P.,
RA Sampson J.R., Jansen B., Hesselting-Janssen A.L.W.,
RA van den Ouweland A.M.W., Eussen B., Verhoef S., Lindhout D.,
RA Halley D.J.J.;
RT "The polycystic kidney disease 1 gene encodes a 14 kb transcript and
RT lies within a duplicated region on chromosome 16. The European
RT Polycystic Kidney Disease Consortium [published erratum appears in
RT Cell 1994 Aug 26;78(4):following 724].";
RT Cell 77:881-894(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95359960; PubMed=7633406;
RA Burn T.C., Connors T.D., Dackowski W.R., Petry L.R., Van Raay T.J.,
RA Milholland J., Venet M., Miller G., Hakim R.M., Doggett N.A.,
RA Landes G.W., Klingner K.W., Qian F., Onuchic L.F., Walnick T.,
RA Germino G.G.;
RT "Analysis of the genomic sequence for the autosomal dominant
RT polycystic kidney disease (PKD) gene predicts the presence of a
RT leucine-rich repeat. The American PKD Consortium (APKD)
RT Consortium";
RT Hum. Mol. Genet. 4:575-582(1995).
DR EMU1; L39891; AAB59488.1;
DR HSBP; P98161; IBA8.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR001024; LH2.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR000483; LRR_Cterm.
DR InterPro: IPR000372; LRR_Nterm.
DR InterPro: IPR003591; LRR_Typ.
DR InterPro: IPR000434; PKD_1.
DR InterPro: IPR000203; PKD_cys_rich.
DR InterPro: IPR000601; PKD_domain.
DR InterPro: IPR002859; REJ.
DR Pfam; PF01825; GCS; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01462; LRRNT; 1.
DR Pfam; PF00801; PKD; 1.
DR Pfam; PF01477; PLAT; 1.
DR Pfam; PF02010; REJ; 1.

```

```

DR PRINTS: PR00500; POLYCYSTIN1.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00303; GPS; 1.
DR SMART: SM00308; LH2; 1.
DR SMART: SM00082; LRRCT; 1.
DR SMART: SM00013; LRRNT; 1.
DR SMART: SM00369; LRR_Typ; 1.
DR SMART: SM00089; PKD; 1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
DR PROSITE: PS50093; PKD; 1.
DR SEQUENCE 3638 AA; 392178 MW; 8267F06B5E04B0D1 CRC64;

```

```

Query Match 12.6%; Score 71.5; DB 4; Length 3638;
Best Local Similarity 26.0%; Pred. No. 88;
Matches 34; Conservative 17; Mismatches 51; Indels 29; Gaps 5;

OY 3 GSLHFAIILAGVDP-----LSMDLPEPRSRASKIRV--HSRCKLWAIHFHGKKSLE 53
Db 2469 GTTAVHGIMLYGVDSRSGRHLDGDRAFHRNSLDIFRIATPHSLGSKVRWHDNKGCL- 2527
OY 54 PSSPPLCTAPHTSLRDOR-----LQSHDLGLGILLKKAIGVSLSRPAPQIO 101
Db 2528 ----SPANFLOHVIVRDQGTARSAFPLYNDWLSVETENGGLVEKVLAA--DALLR 2580
OY 102 YRRLVQIOLK 112
Db 2581 FRRLVAVELOR 2591

```

Search completed: May 30, 2002, 17:25:49  
Job time: 124 sec